

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Bujard, Hermann
Gossen, Manfred
Salfeld, Jochen G.
Voss, Jeffrey W.

(ii) TITLE OF INVENTION: Animals Transgenic for a Tetracycline-Controlled Transcriptional Transactivator

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Lahive & Cockfield
- (B) STREET: 60 State Street
- (C) CITY: Boston
- (D) STATE: Massachusetts
- (E) COUNTRY: USA
- (F) ZIP: 02109-1875

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: ASCII text

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/383,754
(B) FILING DAE: 14-JUN-1994

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/076,327
(B) FILING DAE: 14-JUN-1993

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: DeConti, Giulio A., Jr.
(B) REGISTRATION NUMBER: 31,503
(C) REFERENCE/DOCKET NUMBER: BPI

(iv) TELECOMMUNICATION AUTHORITY

TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 227-7400

(B) TELEFAX: (617) 227-5941

45 INFORMATION FOR SEO ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1008 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Herpes Simplex Virus
- (B) STRAIN: K12, KOS

10

(vii) IMMEDIATE SOURCE

(B) CLONE: tTA transactivator

15

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 1..1008

20

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION: 1..1008

25

(ix) FEATURE:

- (A) NAME/KEY: misc. binding
- (B) LOCATION: 1..207

30

(ix) FEATURE:

- (A) NAME/KEY: misc. binding
- (B) LOCATION: 208..335

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG
Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
1 5 10 15

48

40 CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG
Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
20 25 30

96

45 AAG CTA GGT GTA GAG CAG CCT ACA TTG TAT TGG CAT GTA AAA AAT AAG
Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
35 40 45

144

50 CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT
Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
50 55 60

192

55 ACT CAC TTT TGC CCT TTA GAA GGG GAA AGC TGG CAA GAT TTT TTA CGT
Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg
65 70 75 80

240

55 AAT AAG GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC GAT GGA
Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly
85 90 95

288

	GCA AAA GTA CAT TTA GGT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr 100 105 110	336
5	CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu 115 120 125	384
10	AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys 130 135 140	432
15	GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr 145 150 155 160	480
20	CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu 165 170 175	528
	TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu 180 185 190	576
25	ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCC GCG Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala 195 200 205	624
30	TAC AGC CGC GCG CGT ACG AAA AAC AAT TAC GGG TCT ACC ATC GAG GGC Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly 210 215 220	672
35	CTG CTC GAT CTC CCG GAC GAC GCC CCC GAA GAG GCG GCG GGG CTG GCG Leu Leu Asp Leu Pro Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala 225 230 235 240	720
40	GCT CCG CGC CTG TCC TTT CTC CCC GCG GGA CAC ACG CGC AGA CTG TCG Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser 245 250 255	768
	ACG GCC CCC CCG ACC GAT GTC AGC CTG GGG GAC GAG CTC CAC TTA GAC Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp 260 265 270	816
45	GGC GAG GAC GTG GCG ATG GCG CAT GCC GAC GCG CTA GAC GAT TTC GAT Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp 275 280 285	864
50	CTG GAC ATG TTG GGG GAC GGG GAT TCC CCG GGT CCG GGA TTT ACC CCC Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro 290 295 300	912
55	CAC GAC TCC GCC CCC TAC GGC GCT CTG GAT ATG GCC GAC TTC GAG TTT His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe 305 310 315 320	960
	GAG CAG ATG TTT ACC GAT CCC CTT GGA ATT GAC GAG TAC GGT GGG TAG Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly	1008

325

330

335

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 335 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10	Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu			
	1	5	10	15
	Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln			
		20	25	30
15	Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys			
	35	40	45	
20	Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His			
	50	55	60	
	Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg			
	65	70	75	80
25	Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly			
	85	90	95	
	Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr			
	100	105	110	
30	Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu			
	115	120	125	
35	Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys			
	130	135	140	
	Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr			
	145	150	155	160
40	Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu			
	165	170	175	
	Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu			
	180	185	190	
45	Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala			
	195	200	205	
	Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly			
50	210	215	220	
	Leu Leu Asp Leu Pro Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala			
	225	230	235	240
55	Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser			

245 250 255
Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
260 265 270
5 Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
275 280 285
Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
10 290 295 300
His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
305 310 315 320
15 Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
325 330 335
20

(2) INFORMATION FOR SEQ ID NO:3:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 894 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
25
25 (ii) MOLECULE TYPE: DNA (genomic)
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Herpes Simplex Virus
30 (B) STRAIN: K12, KOS
 (C) INDIVIDUAL ISOLATE: tTAS transactivator
35 (ix) FEATURE:
 (A) NAME/KEY: exon
 (B) LOCATION: 1..894
40
40 (ix) FEATURE:
 (A) NAME/KEY: mRNA
 (B) LOCATION: 1..894
45 (ix) FEATURE:
 (A) NAME/KEY: misc. binding
 (B) LOCATION: 1..207
50 (ix) FEATURE:
 (A) NAME/KEY: misc. binding
 (B) LOCATION: 208..297
50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

55 ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG 48
Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu

	1	5	10	15	
15	CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln				96
	20		25		30
5	AAG CTA GGT GTA GAG CAG CCT ACA TTG TAT TGG CAT GTA AAA AAT AAG Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys				144
	35		40		45
10	CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His				192
	50		55		60
	ACT CAC TTT TGC CCT TTA GAA GGG GAA AGC TGG CAA GAT TTT TTA CGT Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg				240
	65		70		75
15	AAT AAC GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC GAT GGA Asn Asn Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly				288
	85		90		95
	GCA AAA GTA CAT TTA GGT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr				336
	100		105		110
20	CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu				384
	115		120		125
25	AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys				432
	130		135		140
	GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr				480
	145		150		155
30	CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu				528
	165		170		175
	TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu				576
	180		185		190
35	ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCT GAT Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Asp				624
	195		200		205
40	CCA TCG ATA CAC ACG CGC AGA CTG TCG ACG GCC CCC CCG ACC GAT GTC Pro Ser Ile His Thr Arg Arg Leu Ser Thr Ala Pro Pro Thr Asp Val				672
	210		215		220
	AGC CTG GGG GAC GAG CTC CAC TTA GAC GGC GAG GAC GTG GCG ATG GCG Ser Leu Gly Asp Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala				720
	225		230		235
	CAT GCC GAC GCG CTA GAC GAT TTC GAT CTG GAC ATG TTG GGG GAC GGG				768

His Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly
245 250 255

GAT TCC CCG GGT CCG GGA TTT ACC CCC CAC GAC TCC GCC CCC TAC GGC 816
Asp Ser Pro Gly Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly
5 260 265 270

GCT CTG GAT ATG GCC GAC TTC GAG TTT GAG CAG ATG TTT ACC GAT GCC 864
Ala Leu Asp Met Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala
275 280 285

CTT GGA ATT GAC GAG TAC GGT GGG TTC TAG 894
Leu Gly Ile Asp Glu Tyr Gly Phe
10 290 295

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 297 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
20 1 5 . 10 15

Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
25 20 25 30

Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
25 35 40 45

Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
30 50 55 60

Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg
65 70 75 80

Asn Asn Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly
35 85 90 95

Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
100 105 110

Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
40 115 120 125

Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
45 130 135 140

Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
145 150 155 160

Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
50 165 170 175

Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
180 185 190

5 Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Asp
195 200 205

Pro Ser Ile His Thr Arg Arg Leu Ser Thr Ala Pro Pro Thr Asp Val
210 215 220

10 Ser Leu Gly Asp Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala
225 230 235 240

His Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly
15 245 250 255

Asp Ser Pro Gly Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly
260 265 270

20 Ala Leu Asp Met Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala
275 280 285

Leu Gly Ile Asp Glu Tyr Gly Gly Phe
290 295

25

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human cytomegalovirus
(B) STRAIN: K12, Towne

(ix) FEATURE:

- (A) NAME/KEY: mRNA
(B) LOCATION: 382..450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

45 GAATTCCCTCG AGTTTACCAAC TCCCTATCAG TGATAGAGAA AAGTGAAAGT CGAGTTTACC 60
ACTCCCTATC AGTGATAGAG AAAAGTGAAA GTCGAGTTA CCACTCCCTA TCAGTGATAG 120
AGAAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT GAAAGTCGAG 180
TTTACCACTC CCTATCAGTG ATAGAGAAAA GTGAAAGTCG AGTTTACCAAC TCCCTATCAG 240
TGATAGAGAA AAGTGAAAGT CGAGTTTACC ACTCCCTATC AGTGATAGAG AAAAGTGAAA 300
50 GTCGAGCTCG GTACCCGGGT CGAGTAGGCG TGTACGGTGG GAGGCCTATA TAAGCAGAGC 360

TCGTTTAGTG AACCGTCAGA TCGCCTGGAG ACGCCATCCA CGCTGTTTG ACCTCCATAG 420
AAGACACCGG GACCGATCCA GCCTCCGCGG 450

(2) INFORMATION FOR SEQ ID NO:6:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human cytomegalovirus
- (B) STRAIN: Towne

15 (ix) FEATURE: .

- (A) NAME/KEY: mRNA
- (B) LOCATION: 382..450

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAATTCTCG ACCCGGGTAC CGAGCTCGAC TTTCACTTTT CTCTATCACT GATAGGGAGT 60
GGTAAACTCG ACTTTCACTT TTCTCTATCA CTGATAGGGA GTGGTAAACT CGACTTTCAC 120
TTTTCTCTAT CACTGATAGG GAGTGGTAAA CTCGACTTTC ACTTTCTCT ATCACTGATA 180
GGGAGTGGTA AACTCGACTT TCACCTTTCT CTATCACTGA TAGGGAGTGG TAAACTCGAC 240
25 TTTCACTTTT CTCTATCACT GATAGGGAGT GGTAAACTCG ACTTTCACTT TTCTCTATCA 300
CTGATAGGGA GTGGTAAACT CGAGTAGGCG TGTACGGTGG GAGGCCTATA TAAGCAGAGC 360
TCGTTTAGTG AACCGTCAGA TCGCCTGGAG ACGCCATCCA CGCTGTTTG ACCTCCATAG 420
AAGACACCGG GACCGATCCA GCCTCCGCGG 450

(2) INFORMATION FOR SEQ ID NO:7:

30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 398 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Herpes Simplex Virus
- (B) STRAIN: KOS

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGCTCGACT TTCACCTTTC TCTATCACTG ATAGGGAGTG GTAAACTCGA CTTTCACTTT 60

TCTCTATCAC	TGATAGGGAG	TGGTAAACTC	GACTTCACT	TTTCTCTATC	ACTGATAGGG	120	
AGTGGTAAAC	TCGACTTTCA	CTTTCTCTA	TCACTGATAG	GGAGTGGTAA	ACTCGACTTT	180	
CACTTTCTC	TATCACTGAT	AGGGAGTGTT	AAACTCGACT	TTCACTTTTC	TCTATCACTG	240	
ATAGGGAGTG	GTAAAACTCGA	CTTTCACTTT	TCTCTATCAC	TGATAGGGAG	TGGTAAACTC	300	
5	GAGATCCGGC	GAATTGAAAC	ACGCAGATGC	AGTCGGGGCG	GCGCGGTCCG	AGGTCCACTT	360
	CGCATATTAA	GGTGACGCGT	GTGGCCTCGA	ACACCGAG			398

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6244 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

15 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human cytomegalovirus
- (B) STRAIN: Towne (hCMV)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pUHD BGR3

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTCGAGTTA	CCACTCCCTA	TCAGTGATAG	AGAAAAAGTGA	AAGTCGAGTT	TACCACTCCC	60	
TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	TTTACCACTC	CCTATCAGTG	ATAGAGAAAA	120	
GTGAAAGTCG	AGTTTACAC	TCCCTATCAG	TGATAGAGAA	AAAGTGAAGT	CGAGTTTACC	180	
ACTCCCTATC	AGTGATAGAG	AAAAGTGAAG	GTCGAGTTA	CCACTCCCTA	TCAGTGATAG	240	
25	AGAAAAAGTGA	AAGTCGAGTT	TACCACTCCC	TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	300
	CTCGGTACCC	GGGTCGAGTA	GGCGTGTACG	GTGGGAGGCC	TATATAAGCA	GAGCTCGTT	360
	AGTGAACCGT	CAGATCGCCT	GGAGACGCCA	TCCACGCTGT	TTTGACCTCC	ATAGAAAGACA	420
	CCGGGACCGA	TCCAGCCTCC	CGGGCCCCGA	ATTCGAGCTC	GGTACCGGGC	CCCCCCTCGA	480
	GGTCGACGGT	ATCGATAAGC	TTGATATCGA	ATTCCAGGAG	GTGGAGATCC	GCGGGTCCAG	540
30	CCAAACCCA	CACCCATTAA	CTCCTCCCTC	TGCCCTATA	TCCCGGCACC	CCCTCCTCCT	600
	AGCCCTTCC	CTCCTCCCGA	GAGACGGGGG	AGGAGAAAAG	GGGAGTTCAAG	GTCGACATGA	660
	CTGAGCTGAA	GGCAAAGGAA	CCTCGGGCTC	CCCACGTGGC	GGGGGGCGCG	CCCTCCCCCA	720
	CCGAGGTGG	ATCCCAGCTC	CTGGGTGGCC	CGGACCCCTGG	CCCCTCCAG	GGGAGCCAGA	780
	CCTCAGAGGC	CTCGTCTGTA	GTCTCCGCCA	TCCCCATCTC	CCTGGACGGG	TTGCTCTTCC	840

	CCCGGCCCTG TCAGGGGCAG AACCCCCAG ACGGAAGAC GCAGGACCCA CCGTCGTTGT	900
	CAGACGTGGA GGGCGCATT CCTGGAGTCG AAGCCCCGA GGGGCAGGA GACAGCAGCT	960
	CGAGACCTCC AGAAAAGGAC AGCGGCCTGC TGGACAGTGT CCTCGACACG CTCCTGGCGC	1020
	CCTCGGGTCC CGGGCAGAGC CACGCCAGCC CTGCCACCTG CGAGGCCATC AGCCC GTGGT	1080
5	GCCTGTTGG CCCGACCTT CCCGAAGACC CCCGGCTGC CCCCGCTACC AAAGGGGTGT	1140
	TGGCCCCGCT CATGAGCCGA CCCGAGGACA AGGCAGGCAG CAGCTCTGGG ACGGCAGCGG	1200
	CCCACAAAGGT GCTGCCAGG GGACTGTCAC CATCCAGGCA GCTGCTGCTC CCCTCCTCTG	1260
	GGAGCCCTCA CTGGCCGGCA GTGAAGCCAT CCCGCAGCC CGCTGCGGTG CAGGTAGACG	1320
	AGGAGGACAG CTCCGAATCC GAGGGCACCG TGGGCCGCT CCTGAAGGGC CAACCTCGGG	1380
10	CACTGGGAGG CACGGCGGCC GGAGGAGGAG CTGCCCGT CGCGTCTGGA GCGGCCGAG	1440
	GAGGCCTCGC CCTTGTCCCC AAGGAAGATT CTCGCTTCTC GGCCCCCAGG GTCTCCTTGG	1500
	CGGAGCAGGA CGGCCGGTG GCGCCTGGC GCTCCCCGCT GGCCACCTCG GTGGTGGATT	1560
	TCATCCACGT GCCCATCCTG CCTCTCAACC ACGCTTCCT GGCCACCCGC ACCAGGCAGC	1620
	TGCTGGAGGG GGAGAGCTAC GACGGGGGG CCGCGGCCGC CAGCCCCCTC GTCCCGCAGC	1680
15	GGGGCTCCCC CTCTGCCTCG TCCACCCCTG TGGCGGGCGG CGACTTCCCC GACTGCACCT	1740
	ACCCGCCGA CGCCGAGCCC AAAGATGACG CGTTCCCCCT CTACGGCGAC TTCCAGCCGC	1800
	CCGCCCTCAA GATAAAGGAG GAGGAAGAAG CCGCCGAGGC CGCGCGCGC TCCCCGCGTA	1860
	CGTACCTGGT GGCTGGTGCA AACCCGCCG CCTTCCGGGA CTTCCAGCTG GCAGCGCCGC	1920
	CGCCACCCCTC GCTGCCGCCT CGAGTGCCCT CGTCCAGACC CGGGGAAGCG GCGGTGGCGG	1980
20	CCTCCCCAGG CAGTGCCTCC GTCTCCTCCT CGTCCTCGTC GGGTCAACC CTGGAGTGCA	2040
	TCCTGTACAA GGCAGAAGGC GCGCCGCCCC AGCAGGGCCC CTTCGCGCCG CTGCCCTGCA	2100
	AGCCTCCGGG CGCCGGCGCC TGCCTGCTCC CGCGGGACGG CCTGCCCTCC ACCTCCGCT	2160
	CGGGCGCAGC CGCCGGGGCC GCCCCTGCGC TCTACCCGAC GCTCGGCCTC AACGGACTCC	2220
	CGCAACTCGG CTACCAGGCC GCCGTGCTCA AGGAGGGCCT GCCGCAGGTC TACACGCCCT	2280
25	ATCTCAACTA CCTGAGGCCG GATTAGAAG CCAGTCAGAG CCCACAGTAC AGCTTCGAGT	2340
	CACTACCTCA GAAGATTTGT TTGATCTGTG GGGATGAAGC ATCAGGCTGT CATTATGGTG	2400
	TCCTCACCTG TGGGAGCTGT AAGGTCTTCT TTAAAAGGGC AATGGAAGGG CAGCATAACT	2460
	ATTTATGTGC TGGAAGAAAT GACTGCATTG TTGATAAAAT CCGCAGGAAA AACTGCCCGG	2520
	CGTGTGCGCT TAGAAAGTGC TGTCAAGCTG GCATGGCCT TGGAGGGCGA AAGTTAAAA	2580
30	AGTTCAATAA AGTCAGAGTC ATGAGAGCAC TCGATGCTGT TGCTCTCCA CAGCCAGTGG	2640

	GCATTCCAAA TGAAAGCCAA CGAATCACTT TTTCTCCAAG TCAAGAGATA CAGTTAATT	2700
	CCCCCTCTAAT CAACCTGTTA ATGAGCATTG AACCAGATGT GATCTATGCA GGACATGACA	2760
	ACACAAAGCC TGATACCTCC AGTTCTTGC TGACGAGTCT TAATCAACTA GGCGAGCGGC	2820
	AACTTCCTTC AGTGGTAAAAA TGGTCCAAAT CTCTTCCAGG TTTTCGAAAC TTACATATTG	2880
5	ATGACCAGAT AACTCTCATC CAGTATTCTT GGATGAGTTT AATGGTATTT GGACTAGGAT	2940
	GGAGATCCTA CAAACATGTC AGTGGGCAGA TGCTGTATTT TGACACCTGAT CTAATATTAA	3000
	ATGAACAGCG GATGAAAGAA TCATCATTCT ATTCACTATG CCTTACCATG TGGCAGATAAC	3060
	CGCAGGAGTT TGTCAAGCTT CAAGTTAGCC AAGAAGAGTT CCTCTGCATG AAAGTATTAC	3120
	TACTTCTTAA TACAATTCCCT TTGGAAGGAC TAAGAAGTCA AAGCCAGTTT GAAGAGATGA	3180
10	GATCAAGCTA CATTAGAGAG CTCATCAAGG CAATTGGTTT GAGGCAAAAA GGAGTTGTTT	3240
	CCAGCTCACA GCGTTTCTAT CAGCTCACAA AACTTCTTGA TAACTTGCAT GATCTTGTC	3300
	AACAACTTCA CCTGTACTGC CTGAATACAT TTATCCAGTC CCAGGGCGCTG AGTGTGAAT	3360
	TTCCAGAAAT GATGTCTGAA GTTATTGCTG CACAGTTACC CAAGATATTG GCAGGGATGG	3420
	TGAAACCACT TCTCTTTCAT AAAAAGTGAA TGTCAATTAT TTTTCAAAGA ATTAAGTGT	3480
15	GTGGTATGTC TTTCGTTTTG GTCAGGATTA TGACGTCTCG AGTTTTATA ATATTCTGAA	3540
	AGGGAATTCC TGCAGCCCGG GGGATCCACT AGTTCTAGAG GATCCAGACA TGATAAGATA	3600
	CATTGATGAG TTTGGACAAA CCACAACTAG AATGCAGTGA AAAAATGCT TTATTGTGA	3660
	AATTTGTGAT GCTATTGCTT TATTTGTAAC CATTATAAGC TGCAATAAAC AAGTTAACAA	3720
	CAACAATTGC ATTCACTTTA TGTTTCAGGT TCAGGGGGAG GTGTGGGAGG TTTTTAAAG	3780
20	CAAGTAAAAC CTCTACAAAT GTGGTATGGC TGATTATGAT CCTGCAAGCC TCGTCGTCTG	3840
	GCCGGACAC GCTATCTGTG CAAGGTCCCC GGACGCGCGC TCCATGAGCA GAGCGCCCGC	3900
	CGCCGAGGCA AGACTCGGGC GGCGCCCTGC CCGTCCCACC AGGTCAACAG CGGGTAACCG	3960
	GCCTCTTCAT CGGGAATGCG CGCGACCTTC AGCATGCCG GCATGTCCCC TGGCGGACGG	4020
	GAAGTATCAG CTCGACCAAG CTTGGCGAGA TTTTCAGGAG CTAAGGAAGC TAAAATGGAG	4080
25	AAAAAAATCA CTGGATATAC CACCGTTGAT ATATCCAAT GGCATCGTAA AGAACATT	4140
	GAGGCATTTC AGTCAGTTGC TCAATGTACC TATAACCAGA CCGTTCAGCT GCATTAATGA	4200
	ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTCCGC TTCCTCGCTC	4260
	ACTGACTCGC TCGCGCTCGGT CGTTCGCCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG	4320
	GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC	4380
30	CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTCCA TAGGCTCCGC	4440

	CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA	4500
	CTATAAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC	4560
	CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA	4620
	TGCTCACGCT GTAGGTATCT CAGTCGGTG TAGTCGTTG GCTCCAAGCT GGGCTGTGTG	4680
5	CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC	4740
	AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA	4800
	GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT	4860
	AGAAGGACAG TATTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT	4920
	GGTAGCTCTT GATCCGGCAA ACAAACCAACC GCTGGTAGCG GTGGTTTTTT TGTTGCAAG	4980
10	CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG	5040
	TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATT TGTCATGAG ATTATCAAAA	5100
	AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA	5160
	TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG	5220
	ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA	5280
15	CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG	5340
	GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT	5400
	GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT	5460
	TCGCCAGTTA ATAGTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC	5520
	TCGTCGTTG GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA	5580
20	TCCCCCATGT TGTGAAAAAG AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAAGT	5640
	AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC	5700
	ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA	5760
	TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGCGTCAA TACGGATAA TACCGCGCCA	5820
	CATAGCAGAA CTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGCG AAAACTCTCA	5880
25	AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT	5940
	TCAGCATCTT TTACTTTCAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG GCAAATGCC	6000
	GCAAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA	6060
	TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT	6120
	TAGAAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC	6180
30	TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTTT	6240

CGTC

6244

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4963 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

10 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human cytomegalovirus

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pUHD BGR4

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCGAGTTA CCACTCCCTA TCAGTGATAG AGAAAAGTGA AAGTCGAGTT TACCACTCCC	60
TATCAGTGAT AGAGAAAAGT GAAAGTCGAG TTTACCACTC CCTATCAGTG ATAGAGAAAA	120
GTGAAAGTCG AGTTTACCACTC TCCCTATCAG TGATAGAGAA AAGTGAAAGT CGAGTTTACC	180
ACTCCCTATC AGTGATAGAG AAAAGTGAAA GTCGAGTTA CCACTCCCTA TCAGTGATAG	240
20 AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT GAAAGTCGAG	300
CTCGGTACCC GGGTCGAGTA GGC GTGTACG GTGGGAGGCC TATATAAGCA GAGCTCGTT	360
AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC ATAGAACACA	420
CCGGGACCGA TCCAGCCTCC GCGGCCCGA ATTCCGGCCA CGACCATGAC CATGACCCTC	480
CACACCAAAG CATCTGGGAT GGCCCTACTG CATCAGATCC AAGGGAACGA GCTGGAGCCC	540
25 CTGAACCGTC CGCAGCTCAA GATCCCCCTG GAGCGGCCCG TGGCGAGGT GTACCTGGAC	600
AGCAGCAAGC CCGCCGTGTA CAACTACCCC GAGGGCGCCG CCTACGAGTT CAACGCCGCG	660
GCCGCCGCCA ACGCGCAGGT CTACGGTCAG ACCGGCCTCC CCTACGGCCC CGGGTCTGAG	720
GCTGCGCGT TCGGCTCCAA CGGCCTGGGG GTTTCCCCC CACTAACAG CGTGTCTCCG	780
AGCCCGCTGA TGCTACTGCA CCCGCCGCCG CAGCTGTCGC CTTTCTGCA GCCCCACGGC	840
30 CAGCAGGTGC CCTACTACCT GGAGAACGAG CCCAGCGGCT ACACGGTGCG CGAGGCCGGC	900
CCGCCGGCAT TCTACAGGCC AAATTCAAGAT AATCGACGCC AGGGTGGCAG AGAAAGATTG	960
GCCAGTACCA ATGACAAGGG AAGTATGGCT ATGGAATCTG CCAAGGAGAC TCGCTACTGT	1020
GCAGTGTGCA ATGACTATGC TTCAGGCTAC CATTATGGAG TCTGGCCTG TGAGGGCTGC	1080
AAGGCCTTCT TCAAGAGAAG TATTCAAGGA CATAACGACT ATATGTGTCC AGCCACCAAC	1140
35 CAGTGCACCA TTGATAAAAAA CAGGAGGAAG AGCTGCCAGG CCTGCCGGCT CCGCAAATGC	1200

	TACGAAGTGG GAATGATGAA AGGTGGGATA CGAAAAGACC GAAGAGGAGG GAGAATGTTG	1260
	AAACACAAGC GCCAGAGAGA TGATGGGAG GGCAGGGGTG AAGTGGGTC TGCTGGAGAC	1320
	ATGAGAGCTG CCAACCTTG GCCAACCG CTCATGATCA AACGCTCTAA GAAGAACAGC	1380
	CTGGCCTTGT CCCTGACGGC CGACCAGATG GTCATGGCCT TGTTGGATGC TGAGCCCCC	1440
5	ATACTCTATT CCGAGTATGA TCCTACCAGA CCCTTCAGTG AAGCTTCGAT GATGGGCTTA	1500
	CTGACCAACC TGGCAGACAG GGAGCTGGTT CACATGATCA ACTGGCGAA GAGGGTGCCA	1560
	GGCTTTGTGG ATTTGACCCT CCATGATCAG GTCCACCTTC TAGAATGTGC CTGGCTAGAG	1620
	ATCCTGATGA TTGGTCTCGT CTGGCGCTCC ATGGAGCACC CAGTGAAGCT ACTGTTGCT	1680
	CCTAACTTGC TCTTGGACAG GAACCAGGGA AAATGTGTAG AGGGCATGGT GGAGATCTTC	1740
10	GACATGCTGC TGGCTACATC ATCTCGGTTC CGCATGATGA ATCTGCAGGG AGAGGAGTTT	1800
	GTGTGCCTCA AATCTATTAT TTTGCTTAAT TCTGGAGTGT ACACATTTCT CTCCAGCACC	1860
	CTGAAGTCTC TGGAAAGAGAA GGACCATATC CACCGAGTCC TGGACAAGAT CACAGACACT	1920
	TTGATCCACC TGATGGCAA GGCAGGCCTG ACCCTGCAGC AGCAGCACCA GCGGCTGGCC	1980
	CAGCTCCTCC TCATCCTCTC CCACATCAGG CACATGAGTA ACAAAGGCAT GGAGCATCTG	2040
15	TACAGCATGA AGTGCAAGAA CGTGGTCCC CTCTATGACC TGCTGCTGGA GATGCTGGAC	2100
	GCCCACCGCC TACATGCGCC CACTAGCCGT GGAGGGGCAT CCGTGGAGGA GACGGACCAA	2160
	AGCCACTTGG CCACTGCGGG CTCTACTTCA TCGCATTCCCT TGCAAAAGTA TTACATCACG	2220
	GGGGAGGCAG AGGGTTTCCC TGCCACAGTC TGAGAGCTCC CTGGCGGAAT TCGAGCTCGG	2280
	TACCCGGGGA TCCTCTAGAG GATCCAGACA TGATAAGATA CATTGATGAG TTTGGACAAA	2340
20	CCACAACTAG AATGCAGTGA AAAAAATGCT TTATTTGTGA AATTTGTGAT GCTATTGCTT	2400
	TATTTGTAAC CATTATAAGC TGCAATAAAC AAGTTAACAA CAACAATTGC ATTCAATTAA	2460
	TGTTTCAGGT TCAGGGGAG GTGTGGGAGG TTTTTAAAG CAAGTAAACAC CTCTACAAAT	2520
	GTGGTATGGC TGATTATGAT CCTGCAAGCC TCGTCGTCTG GCCGGACCAC GCTATCTGTG	2580
	CAAGGTCCCC GGACGCGCGC TCCATGAGCA GAGCGCCCGC CGCCGAGGCA AGACTCGGGC	2640
25	GGCGCCCTGC CCGTCCCACC AGGTCAACAG GCGGTAACCG GCCTCTTCAT CGGAAATGCG	2700
	CGCGACCTTC AGCATCGCCG GCATGTCCCC TGCGGACGG GAAGTATCAG CTCGACCAAAG	2760
	CTTGGCGAGA TTTTCAGGAG CTAAGGAAGC TAAAATGGAG AAAAAAAATCA CTGGATATAAC	2820
	CACCGTTGAT ATATCCAAT GGCATCGTAA AGAACATTTC GAGGCATTTC AGTCAGTTGC	2880
	TCAATGTACC TATAACCAGA CCGTTCAGCT GCATTAATGA ATCGGCCAAC GCGCGGGGAG	2940
30	AGGCGGTTTG CGTATTGGGC GCTCTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT	3000

	CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA	3060
	ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG	3120
	TAAAAAGGCC GCGTTGCTGG CGTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA	3180
	AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCAGT	3240
5	TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CGGGATAACCT	3300
	GTCCGCCTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT GTAGGTATCT	3360
	CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAAGCC	3420
	CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT	3480
	ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC	3540
10	TACAGAGTTC TTGAAGTGGT GGCTTAACTA CGGCTACACT AGAAGGACAG TATTGGTAT	3600
	CTGCGCTCTG CTGAAGGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA	3660
	ACAAACCACC GCTGGTAGCG GTGGTTTTTG TGTTGCAAG CAGCAGATTA CGCGCAGAAA	3720
	AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA	3780
	AAACTCACGT TAAGGGATTG TGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT	3840
15	TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA	3900
	CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC	3960
	CATAGTTGCC TGATCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC	4020
	CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA	4080
	AACCAGCCAG CCGGAAGGGC CGAGCGCAGA AGTGGTCCTG CAACTTTATC CGCCTCCATC	4140
20	CAGTCTATTA ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTGCGC	4200
	AACGTTGTTG CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG TATGGCTTCA	4260
	TTCAGCTCCG GTTCCCAACG ATCAAGGCAG GTTACATGAT CCCCCATGTT GTGAAAAAAA	4320
	CGGGTTAGCT CCTTCGGTCC TCCGATCGTT GTCAGAAGTA AGTTGGCCGC AGTGTATCA	4380
	CTCATGGTTA TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCCGT AAGATGCTTT	4440
25	TCTGTGACTG GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT	4500
	TGCTCTTGGC CGGCGTCAAT ACGGGATAAT ACCCGGCCAC ATAGCAGAAC TTTAAAAGTG	4560
	CTCATCATTG GAAAACGTT TCAGGGCGA AAACTCTCAA GGATCTTACCGCTGTTGAGA	4620
	TCCAGTTCGA TGTAACCCAC TCGTGCACCC AACTGATCTT CAGCAGCTTT TACTTTCAAC	4680
	AGCGTTCTG GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAAGGG AATAAGGGCG	4740
30	ACACGGAAAT GTTGAATACT CATACTCTTC CTTTTCAAT ATTATTGAAG CATTTATCAG	4800

GGTTATTGTC TCATGAGCGG ATACATATT GAATGTATTT AGAAAAATAA ACAAAATAGGG	4860
GTTCCCGCGCA CATTCCCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT TATTATCATG	4920
ACATTAACCT ATAAAAATAG GCGTATCACG AGGCCCTTTC GTC	4963

5 (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
10 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TCGAGTTTAC CACTCCCTAT CAGTGATAGA GAAAAGTGAA AG

42